

Monge-Kantorovich norms on spaces of vector measures

Ion Chițescu ^a, Radu Miculescu ^{b,*}, Lucian Niță ^c, Loredana Ioana ^d

^a *Faculty of Mathematics and Computer Science, University of Bucharest,
Academiei Str. 14, 010014, Bucharest, Romania*

^b *Faculty of Mathematics and Computer Science, University of Bucharest,
Academiei Str. 14, 010014, Bucharest, Romania*

^c *Technical University of Civil Engineering, Lacul Tei Blvd., 122-124,
020396, Bucharest, Romania*

^d *Faculty of Mathematics and Computer Science, University of Bucharest,
Academiei Str. 14, 010014, Bucharest,
Romania*

Abstract

One considers Hilbert space valued measures on the Borel sets of a compact metric space. A natural numerical valued integral of vector valued continuous functions with respect to vector valued functions is defined. Using this integral, different norms (we called them Monge-Kantorovich norm, modified Monge-Kantorovich norm and Hanin norm) on the space of measures are introduced, generalizing the theory of (weak) convergence for probability measures on metric spaces. These norms introduce new (equivalent) metrics on the initial compact metric space.

MSC 2010: Primary: 28B05, 46G10, 46E10, 28C15. Secondary: 46B25, 46C05

Keywords: Hilbert space; vector measure; vector integral; Lipschitz vector function; weak and strong convergence of measures

* Corresponding author

E-mail addresses: ionchitescu@yahoo.com (Ion Chițescu), miculescu@yahoo.com (Radu Miculescu), luci6691@yahoo.com (Lucian Niță), loredana.madalina.ioana@gmail.com (Loredana Ioana)

1. Introduction

We introduce and study several metrics on certain spaces of vector measures. This is done using a vector integral (with numerical values) previously introduced by us.

A short history of the problem follows. The story began long ago, with the problem of mass transport initiated by G. Monge in 1781 (how to fill up a hole with the material from a given pile of sand, in an optimal way, i.e. with a minimal cost, see [18]). The problem was, actually, very difficult and G. Monge proposed a complicated geometrical solution. Many years after, in 1887, P. Appell completely solved the problem with complicated variational methods (see [1]). L. V. Kantorovich, inventor of linear programming, attacked the problem in a totally different way. First he considered the discreet variant of the problem, "embedding" it in the theory of linear programming and totally solving it (see [12]) in a way suitable for successful use of computers (these results constitute a major part of the reasons for the Nobel prize - of course for economy - received by L.V. Kantorovich). Afterwards, he transformed the problem in an abstract way, working for a compact metric space instead of a finite set and for a measure instead of a vector in \mathbb{R}^n . Alone or jointly with his student G.S. Rubinstein (see [14] and [15]), he succeeded in completely solving the new, abstract problem. The necessary mathematical tools were the theory of normed spaces and different metrics on spaces of measures, let us call them Kantorovich-Rubinstein metrics. In the treatises [13] and [21] these facts are clearly explained, with many details.

The study of different metrics on spaces of measures is closely related to the theory of convergence of probability measures (especially weak convergence) on metric spaces, where the Lipschitz functions play an important role (see [2], [6] and [19]). It is within the framework of this theory that the formalism of Kantorovich-Rubinstein-type metrics appears more clearly.

Our main goal in the present paper is to extend the theories of metrics and convergence in spaces of probabilities (or scalar measures) to (similar) theories in spaces of vector measures. The most suitable framework seemed to us to be the framework of Hilbert space valued measures. So, let X be a Hilbert space.

We needed first an integral. Consequently, we elaborated the theory of a numerical valued integral of continuous functions on a compact metric space taking values in X , with respect to a measure of bounded variation taking values in X . Our integral is sesquilinear (not bilinear in the complex case)

and uniform. In the second part of the paragraph dedicated to preliminaries we expose (without proofs) the main properties of this (natural) integral, among them being some computing devices and an antilinear (not linear in the complex case) and isometric isomorphism between the space of measures and the dual of the space of continuous functions. Details and proofs will appear in "Sesquilinear uniform vector integral", Proceedings - Mathematical Sciences.

Having this integral, we pass in the main paragraph ("Results") of the paper to the study of the space of X -valued measures (of bounded variation) defined on the Borel sets of a compact metric space (T, d) . Different metrics (and locally convex topologies) are introduced on these spaces, a major role being played by the Lipschitz functions (the use of these functions makes the new introduced metrics to be "topologically sensitive", as one can see in the last subparagraph).

We begin with the Monge-Kantorovich norm and the corresponding metric. We continue with the weak* topology. It is seen that, in case of finite dimensional X , the weak* and the Monge-Kantorovich topologies coincide, which is not the case for infinite dimensional X , as one can see later. Afterwards, we introduce a new norm on a subspace of the space of X -valued measures and the corresponding metric on some subsets of the whole space of X -valued measures. We called them "the modified Monge-Kantorovich norm", respectively "the modified Monge-Kantorovich metric". The modified Monge-Kantorovich norm and Monge-Kantorovich norm are equivalent. The results described up to now constitute generalizations (for measures) of many results concerning probability measures.

The next subparagraph generalizes (for vector measures) the ideas of L. Hanin (see [11] and [10]), the ideas being to "extend" the modified Monge-Kantorovich norm to the whole space of X -valued measures (actually one obtains a norm which is equivalent to the modified Monge-Kantorovich norm on the initial subspace). Using the previous results, we introduce a counterexample showing that weak* topology and Monge-Kantorovich topology are not the same for infinite dimensional spaces.

The last subparagraph has a somewhat different character. Namely, using the previously introduced metrics on the space of measures, we can equip the underlying compact metric space T with the corresponding new metrics and we show that all these new metrics are equivalent to the initial metric d (which is not the case of the variational metric).

We feel obliged to add that the use of "Monge-Kantorovich" name which

we preferred seemed suitable to us for historical reasons. Many times the "Kantorovich-Rubinstein" name would have been more correct.

The authors hope to use the results in the present paper in a subsequent paper dedicated to applications (e.g. fractals).

2. Preliminary facts

Notations and general notions

We begin with some notations and general notions appearing throughout the paper.

As usual $\mathbb{N} = \{1, 2, \dots, n, \dots\}$ is the non null positive integers, K is the scalar field (either $K = \mathbb{R}$ or $K = \mathbb{C}$), $\mathbb{R}_+ = \{x \in \mathbb{R} \mid x \geq 0\}$, $\overline{\mathbb{R}}_+ = \mathbb{R}_+ \cup \{\infty\}$, $K^n = \{(x_1, x_2, \dots, x_n) \mid x_i \in K\}$, where $n \in \mathbb{N}$.

For an arbitrary set T , we write $\mathcal{P}(T) = \{A \mid A \subset T\}$ and for $A \subset T$, $\varphi_A : T \rightarrow K$ will be the characteristic (indicator) function of A acting via $\varphi_A(t) = 0$ if $t \notin A$ and $\varphi_A(t) = 1$ if $t \in A$. If $A \subset T$, the complementary set of A is $C_D = \{t \in T \mid t \notin A\}$. We write $(a_i)_{i \in I} \subset T$ (or $(a_i)_i \subset T$) to denote the fact that the family $(a_i)_{i \in I}$ has the property $a_i \in T$ for any $i \in I$. In particular, one can consider sequences $(a_n)_{n \in \mathbb{N}} \subset T$ (or $(a_n)_n \subset T$). In this case, when we write $(a_{n_p})_p \subset (a_n)_n$, this means that $(a_{n_p})_p$ is a subsequence of $(a_n)_n$.

If $f : X \rightarrow Y$ is a function, we shall often write $x \mapsto f(x)$ to designate the fact that the image of $x \in X$ under f is $f(x)$. Assuming f is injective, the (generalized) inverse of f is the function $f^{-1} : f(X) \rightarrow X$ acting via $f^{-1}(y) \stackrel{\text{def}}{=} x$, where $x \in X$ is uniquely determined by the condition $f(x) = y$. Let X, Y, Z be three sets and $A \subset X$, $B \subset Y$. Let $f : A \rightarrow Y$ be such that $f(A) \subset B$ and let $g : B \rightarrow Z$. The (generalized) composition of f and g is the function $g \circ f : A \rightarrow Z$ acting via $(g \circ f)(x) = g(f(x))$.

Let X be a vector space over K . For any $x \in X$, we write $Sp(x)$ for the vector space generated by x , i.e. for $\{\alpha x \mid \alpha \in K\}$. If $f : T \rightarrow K$ is a function, we can define the function $fx : T \rightarrow K$ acting via $fx(t) = f(t)x$.

Now, let us consider a topological space (T, τ) . If $(a_n)_{n \in \mathbb{N}} \subset T$ and $a \in T$, we write $a_n \xrightarrow[n]{} a$ to designate the fact that the sequence $(a_n)_n$ converges to a . Supplementarily, let us consider (more generally) a preordered set (Δ, \leq) ($u \leq v$, for any $u \in \Delta$; $u \leq v$ and $v \leq w$ implies $u \leq w$, for any $u, v, w \in \Delta$) which is directed (for any u, v in Δ there exists $w \in \Delta$ such that $u \leq w$ and $v \leq w$). We consider a function $f : \Delta \rightarrow T$, write $f(\delta) = x_\delta$ for any $\delta \in \Delta$

and identify $f \equiv (x_\delta)_{\delta \in \Delta}$. Under these circumstances, we write $(x_\delta)_{\delta \in \Delta}$ net T (or $(x_\delta)_\delta$ net T). Let also $x \in T$. Then we write $x_\delta \xrightarrow[\delta]{} x$ (and we say that $(x_\delta)_\delta$ converges to x) if for any (basic) neighborhood V of x there exists $\delta(V) \in \Delta$ such that $x_\delta \in V$ whenever $\delta \in \Delta$, $\delta \geq \delta(V)$. For any $A \subset T$ and any $a \in T$, we have the equivalence: $a \in \overline{A}$ (the closure of A) if and only if there exists $(a_\delta)_{\delta \in \Delta}$ net A such that $a_\delta \xrightarrow[\delta]{} a$.

If (T, d) is a metric space, $\emptyset \neq A \subset T$ and $x \in T$, the distance from x to A is defined via $d_1(x, A) = \inf\{d(x, a) \mid a \in A\}$ (clearly $d_1(x, \{a\}) = d(x, a)$). Then $d_1(x, A) = 0$ if and only if $x \in \overline{A}$. For x and y in T one has $|d_1(x, A) - d_1(y, A)| \leq d(x, y)$. Considering two non empty sets $A \subset T$ and $B \subset T$, the distance between A and B is defined via $\delta(A, B) = \inf\{d(a, b) \mid a \in A, b \in B\}$. Clearly $\delta(\{a\}, B) = d_1(a, B)$. If $\emptyset \neq A \subset T$, the diameter of A is $\text{diam}(A) = \sup\{d(x, y) \mid x, y \in A\}$.

For any normed space $(X, \|\cdot\|)$, the dual of X is $X' = \{V : X \rightarrow K \mid V \text{ is linear and continuous}\}$. Then X' becomes a Banach space, when equipped with the (operator) norm $\|V\|_0 = \sup\{|V(x)| \mid x \in X, \|x\| \leq 1\}$. Usually we write only X (instead of $(X, \|\cdot\|)$) in order to designate a normed space.

If X is a Hilbert space, we shall write $(x \mid y)$ for the scalar product of the elements $x, y \in X$. Hence the scalar product $(\cdot \mid \cdot)$ yields the norm $\|\cdot\|$, acting via $\|x\| = \sqrt{(x \mid x)}$.

The space K^n becomes (canonically) a Hilbert space with the scalar product $(x \mid y) = \sum_{i=1}^n x_i \overline{y_i}$, where $x = (x_1, x_2, \dots, x_n)$ and $y = (y_1, y_2, \dots, y_n)$. Hence $\|x\| \geq |x_i|$ for any $i \in \{1, 2, \dots, n\}$, where $x = (x_1, x_2, \dots, x_n)$. The space of sequences $l^2 = \{x = (x_n)_{n \in \mathbb{N}} \mid x_n \in K, \sum_{n=1}^{\infty} |x_n|^2 < \infty\}$ becomes a Hilbert space with the scalar product $(x \mid y) = \sum_{n=1}^{\infty} x_n \overline{y_n}$, where $x = (x_n)_{n \in \mathbb{N}}$ and $y = (y_n)_{n \in \mathbb{N}}$. Let us note that $\|x\| = (\sum_{n=1}^{\infty} |x_n|^2)^{\frac{1}{2}}$, where $x = (x_n)_{n \in \mathbb{N}}$.

For general topology, see [16] and [7]. For functional analysis, see [8], [13] and [20].

A sesquilinear uniform integral

In this subparagraph, we shall present without proofs the sesquilinear uniform integral which will be used throughout the paper.

Let (T, d) be a compact metric space and X a Hilbert space with scalar product $(\cdot | \cdot)$ and corresponding norm $\|x\| = \sqrt{(x | x)}$. The Borel sets of T will be $\mathcal{B} \subset \mathcal{P}(T)$. The vector space $C(X) = \{f : T \rightarrow X \mid f \text{ is continuous}\}$ is a Banach space with norm $f \mapsto \|f\| = \sup\{\|f(t)\| \mid t \in T\}$. Actually, $C(X)$ is a closed space of the Banach space $B(X) = \{f : T \rightarrow X \mid f \text{ is bounded}\}$ equipped with the norm $f \mapsto \|f\| = \sup\{\|f(t)\| \mid t \in T\}$ (the confusional same notation $\|f\|$ for $f \in C(X)$ and $f \in B(X)$ is justified).

A function $f : T \rightarrow X$ will be called *simple* if it has the form $f = \sum_{i=1}^m \varphi_{A_i} x_i$, with $A_i \in \mathcal{B}$ and $x_i \in X$ (one can always consider that the sets A_i are mutually disjoint and $\bigcup_{i=1}^m A_i = T$). One has $S(X) = \{f : T \rightarrow X \mid f \text{ is simple}\} \subset B(X)$ and $S(X)$ is a vector subspace.

A function $\mu : \mathcal{B} \rightarrow X$ is called a σ -additive measure if $\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$ for any sequence $(A_n)_n \subset \mathcal{B}$ of mutually disjoint sets. For such a μ and $A \in \mathcal{B}$, one can define *the variation of μ over A* as follows. We shall say that a finite family $(A_i)_{i \in \{1, 2, \dots, m\}} \subset \mathcal{B}$ is a *partition of A* if the sets A_i are mutually disjoint and $\bigcup_{i=1}^m A_i = A$. Then the variation of μ over A , denoted by $|\mu|(A)$, is defined via $|\mu|(A) = \sup\{\sum_{i=1}^m \|\mu(A_i)\| \mid (A_i)_{i \in \{1, 2, \dots, m\}} \text{ is a partition of } A\}$ (the supremum is taken for all possible partitions of A). We say that μ is of *bounded variation* if $|\mu|(T) < \infty$.

The vector space (with natural operations) $cabv(X) = \{\mu : \mathcal{B} \rightarrow X \mid \mu \text{ is a } \sigma\text{-additive measure of bounded variation}\}$ becomes a Banach space when equipped with the norm $\mu \mapsto \|\mu\| \stackrel{\text{def}}{=} |\mu|(T)$. Any σ -additive measure $\mu : \mathcal{B} \rightarrow K$ is of bounded variation. The topology on $cabv(X)$ generated by this norm will be called the *variational topology* and will be denoted by $\mathcal{T}(var, X)$. For any $a \in (0, \infty)$, let $B_a(X) = \{\mu \in cabv(X) \mid \|\mu\| \leq a\}$. Then $\mathcal{T}(var, X)$ induces the topology $\mathcal{T}(var, X, a)$ on $B_a(X)$. For a sequence $(\mu_n)_n \subset cabv(X)$ and for $\mu \in cabv(X)$, $\mu_n \xrightarrow[n]{var} \mu$ means that $(\mu_n)_n$ converges to μ in $\mathcal{T}(var, X)$. Notice that $\mu_n \xrightarrow[n]{var} \mu \Rightarrow \mu_n \xrightarrow[n]{u} \mu$, the last symbol denoting uniform convergence. Consequently, $\mu_n \xrightarrow[n]{var} \mu$ implies $\mu_n(A) \xrightarrow[n]{} \mu(A)$ for any $A \in \mathcal{B}$. Notice that, if $\mu \in cabv(K)$ and $x \in X$, then $\mu x \in cabv(X)$ and $\|\mu x\| = \|\mu\| \|x\|$.

In the same way, if $(f_n)_n \subset B(X)$ and $f \in B(X)$, we write $f_n \xrightarrow[n]{u} f$ to

denote the fact that $(f_n)_n$ converges uniformly to f (i.e. $(f_n)_n$ converges to f in the Banach space $B(X)$). The closure of $S(X)$ in $B(X)$ is the space of *totally measurable functions* denoted by $TM(X)$. So $TM(X) \stackrel{\text{def}}{=} \overline{S(X)}$. One has $C(X) \subset TM(X)$.

Now, let $\mu \in cabv(X)$. For any $f = \sum_{i=1}^m \varphi_{A_i} x_i \in S(X)$, the integral of f with respect to μ is $\int f d\mu \stackrel{\text{def}}{=} \sum_{i=1}^m (x_i \mid \mu(A_i))$ (the definition does not depend upon the representation of f). Because $|\int f d\mu| \leq \|\mu\| \|f\|$, the linear and continuous map $U : S(X) \rightarrow K$ given via $U(f) = \int f d\mu$ can be extended by uniform continuity to $V : \overline{S(X)} = TM(X) \rightarrow K$. For any $f \in TM(X)$, we write $V(f) \stackrel{\text{def}}{=} \int f d\mu = \text{the integral of } f \text{ with respect to } \mu$. Hence, for any $f \in TM(X)$, one has $\int f d\mu = \lim_n \int f_n d\mu$, where $(f_n)_n \subset S(X)$ is such that $f_n \xrightarrow[n]{u} f$ (the result does not depend upon the sequence $(f_n)_n$ used). So, our integral is uniform and sesquilinear ($\int (\alpha f + \beta g) d\mu = \alpha \int f d\mu + \beta \int g d\mu$ and $\int f d(\alpha \mu + \beta \nu) = \alpha \int f d\mu + \beta \int f d\nu$ for α and β in K , f and g in $TM(X)$ and μ, ν in $cabv(X)$). Notice that, for any $f \in TM(X)$ and any $\mu \in cabv(X)$ one has $|\int f d\mu| \leq \|\mu\| \|f\|$ and $\int f d\mu$ can be computed for any $f \in C(X) \subset TM(X)$. From now on, we shall discuss about $\int f d\mu$ only for $f \in C(X)$.

We feel obliged to insist upon some computational aspects derived from the fact that the complex Hilbert spaces (with their sesquilinear scalar products) make life a bit more complicated.

Firstly, in the particular case when $X = K$, due to the fact that the scalar product in K is given by the formula $(\alpha \mid \beta) = \alpha \overline{\beta}$, we have for any $f \in C(K)$ and any $\mu \in cabv(K)$: $\int f d\mu$ (with the present definition) $= \int f d\overline{\mu}$ (with the standard definition, where $\overline{\mu} \in cabv(K)$ acts via $\overline{\mu}(A) = \overline{\mu(A)}$ for any $A \in \mathcal{B}$). Extending these considerations and considering an orthonormal basis $(e_i)_{i \in I}$ of X , one can identify any $f \in C(X)$ via $f \equiv (f_i)_{i \in I} \subset C(K)$ and any $\mu \in cabv(X)$ via $\mu \equiv (\mu_i)_{i \in I} \subset cabv(K)$, with the following explanations: a) For any $f \in C(X)$ and any $t \in T$, $f(t) = \sum_i f_i(t) e_i$. b) For any $\mu \in cabv(X)$ and any $A \in \mathcal{B}$, $\mu(A) = \sum_i \mu_i(A) e_i$ (summable families). Then one can prove that $\int f d\mu = \sum_i \int f_i d\mu_i$ (the integral being computed with the present definition).

Secondly, we recall the Riesz-Fréchet representation theorem asserting

the existence of the antilinear and isometric bijection $F : X \rightarrow X'$, given via $F(y) = T_y$, where $T_y(x) = (x | y)$ for any $y \in X$ and any $x \in X$. On the basis of this representation theorem, we interpret a (now) classical result of N. Dinculeanu ([5]) and obtain an antilinear and isometric isomorphism $H : cabv(X) \rightarrow C(X)'$ given via $H(\mu) = V_\mu$, where $V_\mu(f) = \int f d\mu$ for any $\mu \in cabv(X)$ and any $f \in C(X)$.

We add some more computational facts. Let $Y \subset X$ be a closed linear subspace and $\pi_Y : X \rightarrow X$ the orthogonal projection defined by Y ($\pi_Y(y) = y$ for any $y \in Y$). Let $f \in C(X)$ and $\mu \in cabv(X)$. Assume that either $f(T) \subset Y$ or $\mu(B) \subset Y$. Then $\int f d\mu = \int (\pi_Y \circ f) d(\pi_Y \circ \mu)$. Other result for $f \in C(K)$, $\mu \in cabv(K)$, $x, y \in X$: one has $\int (fx) d(\mu y) = (\int f d\mu) \cdot (x | y)$ (in case $x = y$ and $\|x\| = 1$, one has $\int (fx) d(\mu x) = (\int f d\mu)$). Finally, we consider, for any $t \in T$, the Dirac measure concentrated at t , namely $\delta_t : \mathcal{B} \rightarrow K$, $\delta_t(A) = \varphi_A(t)$. Then, for any $x \in X$ and any $t \in T$, $\delta_t x \in cabv(X)$ and $\|\delta_t x\| = \|x\|$. For any $f \in C(X)$, one has $\int f d(\delta_t x) = (f(t) | x)$.

For general measure theory see [9] and [17]. For vector measure and vector integration see [5], [4] and [3].

3. Results

The space $L(X)$

From now on, (T, d) will be a compact metric space such that T has at least two elements and X will be a non null Hilbert space with scalar product $(\cdot | \cdot)$ and corresponding norm $\|\cdot\|$.

Recall that a function $f : T \rightarrow X$ is a *Lipschitz function* if there exists a number $M \in (0, \infty)$ such that $\|f(x) - f(y)\| \leq Md(x, y)$ for any x and y in T . For such f , we define the *Lipschitz constant of f* , denoted by $\|f\|_L$, by $\|f\|_L = \sup\{\frac{\|f(x) - f(y)\|}{d(x, y)} \mid x, y \in T, x \neq y\}$. Notice that $\|f\|_L = \min\{M \mid M \geq 0 \text{ such that } \|f(x) - f(y)\| \leq Md(x, y) \text{ for any } x, y \in T\}$.

The space

$$L(X) = \{f : T \rightarrow X \mid f \text{ is a Lipschitz function}\}$$

is seminormed with the seminorm $f \mapsto \|f\|_L$ (we have $\|f\|_L = 0$ if and only if f is constant). The space $L(X)$ is normed with the norm $f \mapsto \|f\|_{BL}$ given via $\|f\|_{BL} = \|f\| + \|f\|_L$. The unit ball of $L(X)$ is $BL_1(X) = \{f \in L(X) \mid \|f\|_{BL} \leq 1\}$.

Of course $L(X) \subset C(X)$. In case $X = K$, $L(K)$ is dense in $C(K)$. This assertion remains valid for $X = K^n$, namely we have

Theorem 1. *For any $n \in \mathbb{N}$, $L(K^n)$ is dense in $C(K^n)$. More precisely, there exists a sequence $(f^m)_m \subset L(K^n)$ such that the set $\{f^m \mid m \in \mathbb{N}\}$ is dense in $C(K^n)$ (which is separable).*

Sketch of the proof. Let $(g^m)_m \subset L(K)$ be a sequence such that $A = \{g^m \mid m \in \mathbb{N}\}$ is dense in $C(K)$. Then the countable set A^n is dense in $C(K^n)$ and this proves all, in view of the following two facts:

a) We have $A \subset L(K^n)$ (because, if $f = (f_1, f_2, \dots, f_n) \in A^n$ and $x, y \in T$, one has

$$\|f(x) - f(y)\| \leq \sum_{i=1}^n |f_i(x) - f_i(y)| \leq \left(\sum_{i=1}^n \|f_i\|_L \right) d(x, y).$$

b) For any $f = (f_1, f_2, \dots, f_n) \in C(K^n)$, there exist n sequences $(g_i^m)_m \subset A$, $i \in \{1, 2, \dots, n\}$, such that $g_i^m \xrightarrow{u} f_i$, for all i and $g^m \xrightarrow{u} f$, where $g^m = (g_1^m, g_2^m, \dots, g_n^m)$. \square

The Monge-Kantorovich norm

In this subparagraph, we introduce the Monge-Kantorovich norm.

The next result is probably well-known, but we think a careful proof of it is desirable. Besides, some technical parts of the proof will be used later.

Theorem 2 (Lipschitz Urysohn-Type Lemma). *Let $\emptyset \neq H \subset T$, $T \neq D \subset T$ such that H is compact, D is open and $H \subset D$. Then there exist a number $M \in (0, 1)$ and a function $f \in BL_1(\mathbb{R})$ with $\|f\|_{BL} = 1$ such that $0 \leq f(t) \leq M$ for any $t \in T$, $f(t) = M$ for any $t \in H$ and $f(t) = 0$ for any $t \in C_D$.*

Proof. a) One has $C_D \neq \emptyset$ and we shall prove that $\delta(H, C_D) > 0$. Indeed, accepting that $\delta(H, C_D) = 0$, we find the sequences $(x_n)_n \subset H$ and $(y_n)_n \subset C_D$ such that $\lim_{n \rightarrow \infty} d(x_n, y_n) = 0$. Due to the compactness, we find (taking convergent subsequences) $x \in H$ and $y \in C_D$ such that $d(x, y) = 0$, i.e. $x = y \in H \cap C_D$, impossible.

b) For any $t \in T$, one has

$$d_1(t, C_D) + d_1(t, H) \geq \delta(C_D, H) > 0 \tag{1}$$

because, for any sequences $(a_n)_n \subset C_D$ and $(b_n)_n \subset H$, one has

$$d(t, a_n) + d(t, b_n) \geq d(a_n, b_n) \geq \delta(C_D, H).$$

So, one can define $g : T \rightarrow \mathbb{R}_+$, via

$$g(t) = \frac{d_1(t, C_D)}{d_1(t, C_D) + d_1(t, H)}$$

and $0 \leq g(t) \leq 1$ for any $t \in T$, $g(t) = 1$ for any $t \in H$ and $g(t) = 0$ for any $t \in C_D$.

For x and y arbitrarily taken in T , using (1), we have:

$$\begin{aligned} |g(x) - g(y)| &= \frac{|d_1(x, C_D)d_1(y, H) - d_1(x, H)d_1(y, C_D)|}{(d_1(x, C_D) + d_1(x, H))(d_1(y, C_D) + d_1(y, H))} \leq \\ &\leq \frac{|d_1(x, C_D)d_1(y, H) - d_1(y, C_D)d_1(y, H) + d_1(y, C_D)d_1(y, H) - d_1(x, H)d_1(y, C_D)|}{(\delta(C_D, H))^2} \leq \\ &\leq \frac{d_1(y, H) |d_1(x, C_D) - d_1(y, C_D)| + d_1(y, C_D) |d_1(y, H) - d_1(x, H)|}{(\delta(C_D, H))^2} \leq \\ &\leq \frac{(d_1(y, H) + d_1(y, C_D))d(x, y)}{(\delta(C_D, H))^2} \leq \frac{2\text{diam}(T)}{(\delta(C_D, H))^2} d(x, y) \end{aligned}$$

The last inequality is true because H and C_D being compact, we have $d_1(y, H) = d(y, h)$ for some $h \in H$ and $d_1(y, C_D) = d(y, p)$ for some $p \in C_D$ a.s.o.

It follows that g is a Lipschitz function and for any x and y in T one has $|g(x) - g(y)| \leq B d(x, y)$, where (T has at least two points) $B = \frac{2\text{diam}(T)}{(\delta(C_D, H))^2} > 0$.

Hence $\|g\| = 1$, $\|g\|_L \leq B$, g is not constant and consequently $1 + B \geq 1 + \|g\|_L = \|g\|_{BL} > 1$.

Finally, we define

$$f = \frac{1}{\|g\|_{BL}} g \text{ and } M = \frac{1}{\|g\|_{BL}}. \quad \square$$

Lemma 3. *Let $\mu_1 : \mathcal{B} \rightarrow \mathbb{R}_+$ and $\mu_2 : \mathcal{B} \rightarrow \mathbb{R}_+$ be two finite σ -additive measures. We have the equivalence: $\mu_1 = \mu_2 \iff \int f d\mu_1 = \int f d\mu_2$ for any positive function $f \in BL_1(\mathbb{R})$.*

Proof. One must prove the implication " \Leftarrow ".

Because T is a metric space, the measures μ_1 and μ_2 are regular: for any $A \in \mathcal{B}$ one has $\mu_1(A) = \sup \mu_1(H)$ and $\mu_2(A) = \sup \mu_2(H)$, the suprema being computed for all compact subsets $H \subset A$. So, it will suffice to prove that $\mu_1(H) = \mu_2(H)$ for any compact subset $H \subset T$. Take such a compact H .

First Possibility: the only open subset D of T such that $H \subset D$ is $D = T$. Again the regularity of μ_1 and μ_2 says that $\mu_i(H) = \inf \{\mu_i(D) \mid H \subset D \subset T, D \text{ open}\}$, hence $\mu_i(H) = \mu_i(T)$, for any $i \in \{1, 2\}$.

Taking $f : T \rightarrow \mathbb{R}$, $f(t) = 1$ for any $t \in T$, one has $f \in BL_1(\mathbb{R})$ and $\int f d\mu_1 = \int f d\mu_2$ (according to the hypothesis), hence $\mu_1(T) = \mu_2(T)$, i.e. $\mu_1(H) = \mu_2(H)$.

Second Possibility: there exist an open set Δ such that $H \subset \Delta \subset T$, $\Delta \neq T$. We shall prove that $\mu_1(H) \leq \mu_2(H)$ (and, in the same way, $\mu_2(H) \leq \mu_1(H)$), hence $\mu_1(H) = \mu_2(H)$.

Take arbitrarily $\varepsilon > 0$. The regularity of μ_2 yields an open set D_ε such that $H \subset D_\varepsilon$ and $\mu_2(D_\varepsilon) \leq \mu_2(H) + \varepsilon$. Write $D = \Delta \cap D_\varepsilon$, hence D is open, $D \neq \emptyset$ and $\mu_2(D) \leq \mu_2(D_\varepsilon) \leq \mu_2(H) + \varepsilon$. For the couple (H, D) , we construct the function f from Theorem 2. According to the hypothesis we have $\int f d\mu_1 = \int f d\mu_2$ and this implies

$$M\mu_1(H) \leq \int f d\mu_1 = \int f d\mu_2 \leq M\mu_2(D) \leq M(\mu_2(H) + \varepsilon),$$

so $\mu_1(H) \leq \mu_2(H) + \varepsilon$. Because ε is arbitrary, it follows that $\mu_1(H) \leq \mu_2(H)$. \square

Theorem 4. For any $\mu \in cabv(X)$, one has the equivalence: $\mu = 0 \Leftrightarrow \int f d\mu = 0$ for any $f \in L(X)$.

Proof. We must prove the implication " \Leftarrow ".

Case $X = \mathbb{R}$. Write $\mu = \mu_1 - \mu_2$, where $\mu_1, \mu_2 \geq 0$ and $\mu_1, \mu_2 \in cabv(\mathbb{R})$ (Jordan decomposition). Hence, for any $f \in L(\mathbb{R})$, $0 = \int f d\mu = \int f d\mu_1 - \int f d\mu_2$. Using Lemma 3, we get $\mu_1 = \mu_2$, hence $\mu = 0$.

Case $X = \mathbb{C}$. Write $\mu = \mu_1 + i\mu_2$, where μ_1 and μ_2 are in $cabv(\mathbb{R})$. Hence, for any $f \in L(\mathbb{R})$, one has $0 = \int f d\mu = \int f d\mu_1 + i \int f d\mu_2 \Leftrightarrow \int f d\mu_1 = \int f d\mu_2 = 0$. Using the case $X = \mathbb{R}$, we get $\mu_1 = \mu_2 = 0$, hence $\mu = 0$.

General case. Let $(e_i)_{i \in I}$ be an orthonormal basis for X . Write, for any $x \in X$: $x = \sum_{i \in I} a_i e_i$, hence $\|x\|^2 = \sum_{i \in I} |a_i|^2$ (summable families, $a_i = (x \mid e_i)$)

Fourier coefficients). For any $i \in I$, define $H_i : X \rightarrow K$ via $H_i(x) = a_i$. Then $H_i \in X'$ and $\|H_i\|_0 = 1$.

For any $\mu \in cabv(X)$ and any $A \in \mathcal{B}$, one has $\mu(A) = \sum_{i \in I} \mu_i(A) e_i$, where $\mu_i = H_i \circ \mu \in cabv(K)$ are uniquely determined. The fact that all μ_i are σ -additive is obvious, due to the continuity of H_i . It follows that all μ_i are of bounded variation (actually $\|\mu_i\| \leq \|\mu\|$ for any $i \in I$, but we will not use this fact).

Now, take $\mu \in cabv(X)$ such that $\int f d\mu = 0$ for any $f \in L(X)$. To prove that $\mu = 0$ means to prove that all $\mu_i = 0$. Fix $i \in I$ arbitrarily. For any $\varphi \in L(K)$, one has $\varphi e_i \in L(X)$ and $\varphi e_i(T) \subset Sp(e_i)$, $(\mu_i e_i)(\mathcal{B}) \subset Sp(e_i)$. Let $Y_i = Sp(e_i)$. Obviously $\varphi e_i = \pi_{Y_i} \circ (\varphi e_i)$ and $\mu_i e_i = \pi_{Y_i} \circ \mu$. According to the hypothesis:

$$\begin{aligned} 0 &= \int (\varphi e_i) d\mu = \int (\pi_{Y_i} \circ (\varphi e_i)) d\mu = \int (\pi_{Y_i} \circ (\varphi e_i)) d(\pi_{Y_i} \circ \mu) = \\ &= \int (\varphi e_i) d(\mu_i e_i) = \left(\int \varphi d\mu_i \right) \|e_i\|^2 = \int \varphi d\mu_i. \end{aligned}$$

Because $\varphi \in L(K)$ is arbitrary, we get $\mu_i = 0$. \square

Theorem 5. *For any $\mu \in cabv(X)$, define*

$$\|\mu\|_{MK} = \sup \left\{ \left| \int f d\mu \right| \mid f \in BL_1(X) \right\}.$$

Then, the function $\mu \mapsto \|\mu\|_{MK}$ is a norm on $cabv(X)$ and one has

$$\|\mu\|_{MK} \leq \|\mu\|$$

for any $\mu \in cabv(X)$.

Proof. Let $\mu \in cabv(X)$. Then, in view of the antilinear identification $C(X)' \equiv cabv(X)$, one has

$$\|\mu\| = \sup \left\{ \left| \int f d\mu \right| \mid f \in B^1(X) \right\},$$

where $B^1(X) = \{f \in C(X) \mid \|f\| \leq 1\}$ and $\|f\| \leq \|f\|_{BL}$ implies $BL_1(X) \subset B^1(X)$, hence $\|\mu\|_{MK} \leq \|\mu\|$.

It is obvious that $\mu \mapsto \|\mu\|_{MK}$ is a seminorm. To finish the proof, one must show the implication $\|\mu\|_{MK} = 0 \Rightarrow \mu = 0$.

But $\|\mu\|_{MK} = 0$ means $\int f d\mu = 0$ for any $f \in L(X)$ and this implies $\mu = 0$, according to Theorem 4. \square

Definition 6. The norm $\|\cdot\|_{MK}$ is called the *Monge-Kantorovich norm*.

Notice that, according to the definition, one has, for any $\mu \in cabv(X)$ and any $f \in L(X)$:

$$\left| \int f d\mu \right| \leq \|\mu\|_{MK} \|f\|_{BL}. \quad (2)$$

For any $t \in T$ and any $x \in X$, $\|x\| = 1$, one has

$$\|\delta_t x\|_{MK} = 1. \quad (3)$$

Indeed, write $\delta_t x = \mu$ and take $f \in BL_1(X)$. One has $\int f d\mu = (f(t) \mid x)$, hence $|\int f d\mu| \leq \|f(t)\| \|x\| = \|f(t)\| \leq 1$ and $\|\mu\|_{MK} \leq 1$. Conversely, taking $f(t) = x$ for any $t \in T$, one has $\|f\|_{BL} = \|f\| = 1$, hence $|\int f d\mu| = (x \mid x) = 1$ and $\|\mu\|_{MK} \geq 1$ a.s.o.

The topology generated by $\|\cdot\|_{MK}$ on $cabv(X)$ will be denoted by $\mathcal{T}(MK, X)$ (*the Monge-Kantorovich topology*). For any $a > 0$, the topology induced by $\mathcal{T}(MK, X)$ on $B_a(X)$ will be denoted by $\mathcal{T}(MK, X, a)$. For a sequence $(\mu_n)_n \subset cabv(X)$ and for $\mu \in cabv(X)$, we shall write $\mu_n \xrightarrow[n]{MK} \mu$ to denote the fact that $(\mu_n)_n$ converges to μ in the Monge-Kantorovich topology.

In the sequel, we shall make some considerations concerning the comparison between the variational topology $\mathcal{T}(var, X)$ and the Monge-Kantorovich topology $\mathcal{T}(MK, X)$.

Due to the inequality $\|\mu\|_{MK} \leq \|\mu\|$, we have $\mathcal{T}(MK, X) \subset \mathcal{T}(var, X)$. Of course, if T is finite, one has $\mathcal{T}(MK, X) = \mathcal{T}(var, X)$. As concerns the case when T is infinite, we remark first that T is infinite if and only if T has at least an accumulation point. Here comes

Theorem 7. *Assume T is infinite. Then the inclusion $\mathcal{T}(MK, X) \subset \mathcal{T}(var, X)$ is strict. Also in this case, the normed space $(cabv(X), \|\cdot\|_{MK})$ is not Banach.*

Proof. a) First we shall prove that for any a and b in T , $a \neq b$ and any $x \in X$, $\|x\| = 1$, one has

$$\|\delta_a x - \delta_b x\| = 2 \quad (4)$$

(in case $X = K$, $x = 1$, one has $\|\delta_a - \delta_b\| = 2$).

To prove (4), write $\delta_a x - \delta_b x = \mu$ and take a partition $(A_i)_{i \in \{1, 2, \dots, m\}}$ of T . One has either $\sum_{i=1}^m \|\mu(A_i)\| = 0$ (in case there exists i such that $a \in A_i$ and $b \in A_i$) or $\sum_{i=1}^m \|\mu(A_i)\| = 2$ (in case there exist $i \neq j$ such that $a \in A_i$ and $b \in A_j$). The second alternative is always possible, taking $A_1 = B(a, r)$, $A_2 = B(b, r)$ with $A_1 \cap A_2 = \emptyset$ and the other A_i arbitrarily. Hence, by passing to supremum, one gets $\|\mu\| = 2$.

b) Again, for a and b in T , $a \neq b$ and any $x \in X$, $\|x\| = 1$, we shall prove that

$$\|\delta_a x - \delta_b x\|_{MK} \leq d(a, b). \quad (5)$$

Indeed, writing again $\mu = \delta_a x - \delta_b x$, we have, for any $f \in L(X)$: $\int f d\mu = (f(a) - f(b) \mid x)$ (in case $X = K$, $x = 1$: $\int f d\mu = f(a) - f(b)$). Hence, if $f \in BL_1(X)$, one has

$$\left| \int f d\mu \right| \leq \|f(a) - f(b)\| \|x\| = \|f(a) - f(b)\| \leq d(a, b)$$

and passing to supremum, we get (5).

At the end of the paper, we shall discuss supplementarily formula (5).

c) Because T is infinite, we take an accumulation point $t_0 \in T$ and a sequence $(t_n)_n \subset T$ such that $t_n \xrightarrow[n]{} t_0$ and $t_n \neq t_0$ for any $n \geq 1$. According to (5), it follows that, for any $x \in X$ with $\|x\| = 1$, one has $\delta_{t_n} x \xrightarrow[n]{MK} \delta_{t_0} x$, whereas, according to (4), the assertion $\delta_{t_n} x \xrightarrow[n]{var} \delta_{t_0} x$ is false. Hence the inclusion $\mathcal{T}(MK, X) \subset \mathcal{T}(var, X)$ must be strict.

The fact that $(cabv(X), \|\cdot\|_{MK})$ is not Banach follows from the inequality $\|\cdot\|_{MK} \leq \|\cdot\|$ and from the fact that the norms $\|\cdot\|_{MK}$ and $\|\cdot\|$ are not equivalent. \square

Let us present some Supplementary Remarks

Remarks

a) If T is infinite, one can find a sequence $(\mu_n)_n \subset cabv(X)$ such that $\|\mu_n\|_{MK} = 1$ and $\|\mu_n\| > n$ for any n .

b) Generally speaking one has for any sequence $(\mu_n)_n \subset cabv(X)$ the implication $\mu_n \xrightarrow[n]{var} \mu \Rightarrow \mu_n \xrightarrow[n]{MK} \mu$. The converse implication is not true for infinite T .

Example

Take $T = [0, 1]$, $X = \mathbb{R}$ and consider the true fact that $\delta_{\frac{1}{n}} \xrightarrow[n]{\text{MK}} \delta_0$. Because, for any $n \in \mathbb{N}$, one has $\delta_{\frac{1}{n}}((0, 1]) = 1$ and $\delta_0((0, 1]) = 0$, it follows that $\delta_{\frac{1}{n}} \xrightarrow[n]{\text{var}} \delta_0$ is a false assertion because, generally speaking $\mu_n \xrightarrow[n]{\text{var}} \mu$ implies $\mu_n(A) \rightarrow \mu(A)$ for any $A \in \mathcal{B}$.

So, *convergence in the Monge-Kantorovich norm does not imply pointwise convergence*. We shall see later that convergence in the Monge-Kantorovich topology means weak* convergence for bounded sequences and this explains everything.

The weak* topology on $\text{cabv}(X)$

Let us introduce a new topology on $\text{cabv}(X)$. This topology is defined on the basis of the fact that $\text{cabv}(X)$ is identified with the dual of $C(X)$.

Definition 8. *The weak* topology on $\text{cabv}(X)$ is the (separated) locally convex topology on $\text{cabv}(X)$ generated by the family of seminorms $(p_f)_{f \in C(X)}$, where, for any $f \in C(X)$, $p_f : \text{cabv}(X) \rightarrow \mathbb{R}_+$ is given via*

$$p_f(\mu) = \left| \int f d\mu \right|.$$

The weak* topology will be denoted by $\mathcal{T}(w^*, X)$ and, for any $a > 0$, its restriction to $B_a(X)$, will be denoted by $\mathcal{T}(w^*, X, a)$.

For any $\mu \in \text{cabv}(X)$, a neighborhood basis for μ is formed with all sets of the form

$$V(\mu; g_1, g_2, \dots, g_m; \varepsilon) \stackrel{\text{def}}{=} \left\{ v \in \text{cabv}(X) \mid \left| \int g_i d(\mu - v) \right| < \varepsilon, i \in \{1, 2, \dots, m\} \right\}$$

(one takes into consideration all possible $\varepsilon > 0$, all $m \in \mathbb{N}$ and all $g_i \in C(X)$).

For a sequence $(\mu_n)_n \subset \text{cabv}(X)$ and for $\mu \in \text{cabv}(X)$, we shall write $\mu_n \xrightarrow[m]{w^*} \mu$ to denote the fact that $(\mu_n)_n$ converges to μ in $\mathcal{T}(w^*, X)$. This means that $\lim_m \int f d\mu_m = \int f d\mu$ for any $f \in C(X)$.

Notice that Alaoglu's theorem implies that, for any $a > 0$, the set $B_a(X)$ is weak* compact (i.e. compact in $\mathcal{T}(w^*, X)$).

In the sequel, we shall fix $n \in \mathbb{N}$ and we shall work for $X = K^n$. We have seen (Theorem 1) that one can find a sequence $(f_m)_m \subset L(K^n)$ such that $\{f_m \mid m \in \mathbb{N}\}$ is dense in $C(K^n)$. This fact has the following two important consequences.

Theorem 9 (Metrisability of $B_a(K^n)$ under $\mathcal{T}(w^*, K^n)$). *For any $a > 0$, the topology $\mathcal{T}(w^*, K^n, a)$ is metrisable. The set $B_a(K^n)$ is compact as a subset of the topological space $(cabv(K^n), \mathcal{T}(w^*, K^n))$. Consequently, $B_a(K^n)$ considered as a metric space (with any metric generating $\mathcal{T}(w^*, K^n, a)$) is complete.*

The metrisability of $\mathcal{T}(w^*, K^n, a)$ follows from the separability of $C(K^n)$, viewing $cabv(K^n)$ as the dual of $C(K^n)$ (see [8], V, 5.1, page 426).

Theorem 10. *Let $(f_m)_m \subset L(K^n)$ be the aforementioned dense sequence in $C(K^n)$. Then, for any $a > 0$, any sequence $(\mu_p)_p \subset B_a(K^n)$ and any $\mu \in B_a(K^n)$ one has the equivalence: $\mu_p \xrightarrow[p]{w^*} \mu \Leftrightarrow \int f_m d\mu_p \xrightarrow[p]{} \int f_m d\mu$ for any $m \in \mathbb{N}$.*

Proof. Only the implication " \Leftarrow " must be proved. Let us consider $V = V(\mu; g_1, g_2, \dots, g_p; \varepsilon) \cap B_a(X)$ a basic neighborhood of μ in $\mathcal{T}(w^*, X, a)$. For any $m \in \{1, 2, \dots, p\}$, choose f_{i_m} such that $\|f_{i_m} - g_m\|_\infty < \frac{\varepsilon}{4a}$. Take $\delta = \min\{\frac{\varepsilon}{3}, \frac{\varepsilon}{4a}\}$, construct $W = V(\mu; f_{i_1}, f_{i_2}, \dots, f_{i_p}; \delta) \cap B_a(X)$ and notice that $W \subset V$. Indeed, if $v \in W$, one has, for any $m \in \{1, 2, \dots, p\}$:

$$\begin{aligned} \left| \int g_m d(\mu - v) \right| &\leq \left| \int (g_m - f_{i_m}) d(\mu - v) \right| + \left| \int f_{i_m} d(\mu - v) \right| \leq \\ &\leq \|g_m - f_{i_m}\| \|\mu - v\|(T) + \delta \leq \frac{\varepsilon}{4a} \|\mu - v\| + \frac{\varepsilon}{3} \leq \\ &\leq \frac{\varepsilon}{4a} (\|\mu\| + \|v\|) + \frac{\varepsilon}{3} \leq \frac{\varepsilon}{4a} (a + a) + \frac{\varepsilon}{3} < \varepsilon. \end{aligned}$$

Using the hypothesis, one can find $n_V \in \mathbb{N}$ such that $\mu_n \in W \subset V$ for any $n \in \mathbb{N}$, $n \geq n_V$. \square

We shall need:

Theorem 11 (Arzela-Ascoli-Type Theorem). *For any $n \in \mathbb{N}$, the set $BL_1(K^n)$ is relatively compact in $C(K^n)$.*

Proof. Let $(f^m)_m$ be a sequence in $BL_1(K^n)$, with each $f^m = (f_1^m, f_2^m, \dots, f_n^m)$, $f_i^m \in C(K)$. Due to the fact that for $x = (x_1, x_2, \dots, x_n) \in K^n$ one has $\|x\| \geq |x_i|$, we get $\|f_i^m\| \leq \|f^m\| \leq \|f^m\|_{BL} \leq 1$ and $\|f_i^m\|_L \leq \|f^m\|_L \leq \|f^m\|_{BL} \leq 1$ for any $m \in \mathbb{N}$ and $i \in \{1, 2, \dots, n\}$. Hence the sequence $(f_1^m)_m$ is bounded and equicontinuous in $C(K)$. Using the Arzela-Ascoli Theorem, we find a subsequence $(f_1^{m_1^p})_p \subset (f_1^m)_m$ and a function $f_1 \in C(K)$ such that $f_1^{m_1^p} \xrightarrow[p]{u} f_1$. Continuing, we find $(f_2^{m_2^p})_p \subset (f_1^{m_1^p})_p$ and $f_2 \in C(K)$ such that $f_2^{m_2^p} \xrightarrow[p]{u} f_2$ and so on. Finally, we find $(f_n^{m_n^p})_p \subset (f_2^{m_2^p})_p$ and $f_n \in C(K)$ such that $f_n^{m_n^p} \xrightarrow[p]{u} f_n$.

It follows that $(f_n^{m_n^p})_p \subset (f^m)_m$ and $f_n^{m_n^p} \xrightarrow[p]{u} f = (f_1, f_2, \dots, f_n) \in C(K^n)$.

□

Remark. It is natural to ask whether the previous result remains valid for an arbitrary Hilbert space X instead of K^n (i.e. if $BL_1(X)$ is relatively compact in $C(X)$ also for infinite dimensional X). The answer is negative, as we shall see later.

We begin the investigation of the connection between the topologies $\mathcal{T}(w^*, K^n)$ and $\mathcal{T}(MK, K^n)$.

Theorem 12 (Coincidence of weak*-Convergence and Monge-Kantorovich Convergence). *Let $a > 0$ and $n \in \mathbb{N}$. For a sequence $(\mu_m)_m \subset B_a(K^n)$ and for $\mu \in B_a(K^n)$ we have the equivalence: $\mu_m \xrightarrow[m]{MK} \mu \Leftrightarrow \mu_m \xrightarrow[m]{w^*} \mu$.*

Proof. The implication " \Rightarrow " Accept that $\mu_m \xrightarrow[m]{MK} \mu$. In view of Theorem 10, one must prove that, for any $p \in \mathbb{N}$ one has $\lim_m \int f_p d\mu_m = \int f_p d\mu$, where $(f_p)_p \subset L(K^n)$ is a dense sequence in $C(K^n)$.

So take an arbitrary $f_p \neq 0$ and let $g = \frac{1}{\alpha} f_p$, where $\alpha = \|f_p\|_{BL}$, hence $\|g\|_{BL} = 1$. Take also $\varepsilon > 0$ arbitrarily. Our hypothesis being that

$$\lim_m (\sup \{ \left| \int h d(\mu_m - \mu) \right| \mid h \in BL_1(K^n) \}) = 0,$$

one can find $m_\varepsilon \in \mathbb{N}$ such that for any $m \in \mathbb{N}$, $m \geq m_\varepsilon$ and any $h \in BL_1(K^n)$,

one has

$$\left| \int h d(\mu_m - \mu) \right| < \frac{\varepsilon}{\alpha},$$

hence

$$\left| \int g d(\mu_m - \mu) \right| < \frac{\varepsilon}{\alpha},$$

which means

$$\left| \int f_p d(\mu_m - \mu) \right| < \varepsilon.$$

The implication " \Leftarrow " Accept (reductio ad absurdum) the existence of a sequence $(\mu_m)_m \subset B_a(K^n)$ and of a $\mu \in B_a(K^n)$ such that $\mu_m \xrightarrow[m]{w^*} \mu$ and such that the assertion $\mu_m \xrightarrow[m]{MK} \mu$ is false. We shall arrive at a contradiction.

Indeed, there exists $\varepsilon_0 > 0$ and a subsequence $(\mu_{m_p})_{m_p} \subset (\mu_m)_m$ such that $\|\mu_{m_p} - \mu\|_{MK} > 2\varepsilon_0$ for any p . So, for any p , one can find $f_p \in BL_1(K^n)$ such that

$$\left| \int f_p d(\mu_{m_p} - \mu) \right| > \varepsilon_0. \quad (6)$$

Using Theorem 11, one can find $(f_{p_q})_q \subset (f_p)_p$ and $f \in C(K^n)$ such that $\|f_{p_q} - f\|_q \rightarrow 0$. Because $\mu_{m_p} \xrightarrow[p]{w^*} \mu$, we get $p_1 \in \mathbb{N}$ such that, for any $p \geq p_1$, one has

$$\left| \int f d(\mu_{m_p} - \mu) \right| < \frac{\varepsilon_0}{2}. \quad (7)$$

Let $q_1 \in \mathbb{N}$ be such that $p_{q_1} > p_1$, and, for any $q \geq q_1$, one has

$$\|f_{p_q} - f\| < \frac{\varepsilon_0}{4a}. \quad (8)$$

From (6), it follows that, for any $q \geq q_1$, one has

$$\left| \int f_{p_q} d(\mu_{m_{p_q}} - \mu) \right| > \varepsilon_0. \quad (9)$$

At the same time, for such q , one has

$$\left| \int f_{p_q} d(\mu_{m_{p_q}} - \mu) \right| \leq \left| \int (f_{p_q} - f) d(\mu_{m_{p_q}} - \mu) \right| + \left| \int f d(\mu_{m_{p_q}} - \mu) \right| \leq$$

$$\begin{aligned}
&\leq \|f_{p_q} - f\| \|\mu_{m_{p_q}} - \mu\| + \left| \int f d(\mu_{m_{p_q}} - \mu) \right| \leq \\
&\leq \|f_{p_q} - f\| (\|\mu_{m_{p_q}}\| + \|\mu\|) + \left| \int f d(\mu_{m_{p_q}} - \mu) \right| \leq \\
&\leq 2a \|f_{p_q} - f\| + \left| \int f d(\mu_{m_{p_q}} - \mu) \right| < \frac{\varepsilon_0}{4a} 2a + \frac{\varepsilon_0}{2} = \varepsilon_0,
\end{aligned}$$

where we used (7) and (8). This contradicts (9). \square

Let us interpret the last results. Take arbitrarily $a > 0$ and $n \in \mathbb{N}$. On $B_a(K^n)$ we have two metrisable topologies: $\mathcal{T}(MK, K^n, a)$ and $\mathcal{T}(w^*, K^n, a)$ (with Theorem 9). Theorem 12 says that the convergent sequences coincide in these topologies, hence they are equal:

$$\mathcal{T}(MK, K^n, a) = \mathcal{T}(w^*, K^n, a). \quad (10)$$

Again Theorem 9 says that $B_a(K^n)$ is compact for $\mathcal{T}(w^*, K^n, a)$, hence for $\mathcal{T}(MK, K^n, a)$. So $B_a(K^n)$ is a compact (hence complete) metric space for the metric given by $\|\cdot\|_{MK}$.

We got (see Theorem 7 too):

Theorem 13. *For any $a > 0$ and any $n \in \mathbb{N}$, the set $B_a(K^n)$, equipped with the metric generated by the Monge-Kantorovich norm $\|\cdot\|_{MK}$, is a compact, hence complete, metric space, its topology being exactly $\mathcal{T}(w^*, K^n, a)$ (in spite of the fact that the normed space $(cabv(K^n), \|\cdot\|_{MK})$ is not complete if T is infinite).*

Remark. The "basis" of Theorem 13 is Theorem 12 which asserts the coincidence of convergent sequences in $\mathcal{T}(MK, K^n, a)$ and $\mathcal{T}(w^*, K^n, a)$. This coincidence is no longer valid for general X instead of K^n as we shall see later.

The modified Monge-Kantorovich norm

In this subparagraph, we shall be concerned with the so called "modified Monge-Kantorovich norm", which can be defined only on a subspace of $cabv(X)$. This new norm is strongly related to the Monge-Kantorovich norm and generates a most important distance (which generalizes classical Kantorovich-Rubinstein metric on the space of probabilities, see e.g [6]) on some distinguished subsets of $cabv(X)$.

For any $v \in X$, let us define

$$cabv(X, v) = \{\mu \in cabv(X) \mid \mu(T) = v\}.$$

Clearly $\delta_t v \in cabv(X, v)$ for any $t \in T$. It is seen that $cabv(X, 0)$ is a vector subspace of $cabv(X)$. For any $\emptyset \neq A \subset cabv(X, v)$ one has $A - A \stackrel{def}{=} \{\mu - \nu \mid \mu, \nu \in A\} \subset cabv(X, 0)$.

Lemma 14. *For any $v \in X$, the set $cabv(X, v)$ is weak* closed in $cabv(X)$.*

Proof. Take arbitrarily $x \in X$ and let us define the constant function $\varphi_x : T \rightarrow X$, given via $\varphi_x(t) = x$, for any $t \in T$.

Now take an arbitrary adherent point $\mu \in cabv(X)$ for $cabv(X, v)$. Hence one can find $(\mu_\delta)_\delta$ net $cabv(X, v)$ such that $\mu_\delta \xrightarrow{\delta} \mu$ in the topology $\mathcal{T}(w^*, X)$, i.e. $\int f d\mu_\delta \xrightarrow{\delta} \int f d\mu$ for any $f \in C(X)$. Consequently $\int \varphi_x d\mu_\delta \xrightarrow{\delta} \int \varphi_x d\mu$ for any $x \in X$, which means $(x \mid \mu_\delta(T)) = (x \mid v) \xrightarrow{\delta} (x \mid \mu(T))$. So $(x \mid \mu(T)) = (x \mid v)$ for any $x \in X$, hence $\mu(T) = v$ and $\mu \in cabv(X, v)$. \square

Define

$$L_1(X) = \{f \in L(X) \mid \|f\|_L \leq 1\}$$

and clearly $BL_1(X) \subset L_1(X)$.

For any $\mu \in cabv(X, 0)$, let us define

$$\|\mu\|_{MK}^* \stackrel{def}{=} \sup\left\{\left|\int f d\mu\right| \mid f \in L_1(X)\right\}.$$

Theorem 15. *For any $\mu \in cabv(X, 0)$, one has*

$$\|\mu\|_{MK} \leq \|\mu\|_{MK}^* \leq \|\mu\| \operatorname{diam}(T).$$

Proof. The first inequality is given by the inclusion $BL_1(X) \subset L_1(X)$.

To prove the second inequality, let us take arbitrarily $f \in L_1(X)$. For any $t_0 \in T$, one has

$$\begin{aligned} \left|\int f d\mu\right| &= \left|\int (f - f(t_0)) d\mu + \int f(t_0) d\mu\right| = \\ &= \left|\int (f - f(t_0)) d\mu + (f(t_0) \mid \mu(T))\right| = \left|\int (f - f(t_0)) d\mu\right| \leq \end{aligned}$$

$$\leq \|f - f(t_0)\| \|\mu\|.$$

Because $\|f(t) - f(t_0)\| \leq \text{diam}(T)$, for any $t \in T$, one has $\|f - f(t_0)\| \leq \text{diam}(T)$, so $|\int f d\mu| \leq \|\mu\| \text{diam}(T)$. \square

Remark. According to the definition, one has for any $\mu \in \text{cabv}(X, 0)$ and any $f \in L(X)$:

$$\left| \int f d\mu \right| \leq \|\mu\|_{MK}^* \|f\|_L. \quad (10)$$

Indeed, in case $\|f\|_L = 0$, i.e. $f \equiv v \in X$ (f is constant), one has $\int f d\mu = (v \mid \mu(T)) = 0$. In case $\|f\|_L > 0$, take $g = \frac{1}{\|f\|_L} f$ and $g \in L_1(X)$, hence $|\int g d\mu| \leq \|\mu\|_{MK}^*$ a.s.o.

Theorem 16. The function $p : \text{cabv}(X, 0) \rightarrow \mathbb{R}_+$ given via $p(\mu) = \|\mu\|_{MK}^*$ is a norm on $\text{cabv}(X, 0)$.

Proof. Using Theorem 15, one can see that p takes finite values and that $p(\mu) = 0$ if and only if $\mu = 0$. The fact that p is a seminorm is obvious. \square

Definition 17. The norm $\|\cdot\|_{MK}^*$ defined above on $\text{cabv}(X, 0)$ is called the modified Monge-Kantorovich norm.

Theorem 18. The norms $\|\cdot\|_{MK}$ and $\|\cdot\|_{MK}^*$ are equivalent on $\text{cabv}(X, 0)$. More precisely, for any $\mu \in \text{cabv}(X, 0)$, one has

$$\|\mu\|_{MK} \leq \|\mu\|_{MK}^* \leq \|\mu\|_{MK} (\text{diam}(T) + 1)$$

and

$$\|\mu\|_{MK} \leq \|\mu\|_{MK}^* \leq \|\mu\| \text{diam}(T)$$

Proof. Take $\mu \in \text{cabv}(X, 0)$. It remains to be proved that

$$\|\mu\|_{MK}^* \leq \|\mu\|_{MK} (\text{diam}(T) + 1).$$

For arbitrary $f \in L_1(X)$ and $t_0 \in T$, define $h : T \rightarrow K$ via $h(t) = f(t) - f(t_0)$. Then $\|h\|_L = \|f\|_L \leq 1$ and (obviously) $\|h\| \leq \text{diam}(T)$. Consequently $\|h\|_{BL} \leq \text{diam}(T) + 1$. Because $\mu(T) = 0$, one has $\int f d\mu = \int h d\mu$, hence (see (2))

$$\left| \int f d\mu \right| = \left| \int h d\mu \right| \leq \|\mu\|_{MK} \|h\|_{BL} \leq \|\mu\|_{MK} (\text{diam}(T) + 1)$$

and f is arbitrary. \square

Theorem 18 says that the topology $\mathcal{T}(MK^*, X)$ generated by $\|\cdot\|_{MK}^*$ on $cabv(X, 0)$ coincides with the topology induced by $\mathcal{T}(MK, X)$ on $cabv(X, 0)$. For a sequence $(\mu_n)_{n \in \mathbb{N}} \subset cabv(X, 0)$ and for $\mu \in cabv(X, 0)$, one has the equivalence: $\mu_n \xrightarrow[n]{MK^*} \mu$ if and only if $\mu_n \xrightarrow[n]{MK} \mu$.

Theorem 19. *Let a and b be in T , $a \neq b$ and $x \in X$, $\|x\| = 1$. Then $\delta_a x - \delta_b x \in cabv(X, 0)$ and*

$$\|\delta_a x - \delta_b x\|_{MK}^* = d(a, b).$$

Proof. For any $f \in L_1(X)$, writing $\mu = \delta_a x - \delta_b x$, one has:

$$\begin{aligned} \left| \int f d\mu \right| &= |(f(a) | x) - (f(b) | x)| \leq \|f(a) - f(b)\| \|x\| = \\ &= \|f(a) - f(b)\| \leq d(a, b), \end{aligned}$$

hence

$$\|\mu\|_{MK}^* \leq d(a, b).$$

Define $f : T \rightarrow X$, via $f(t) = d(t, a)x$. Then $f \in L_1(X)$, because, for u and v in T , one has:

$$\begin{aligned} \|f(u) - f(v)\| &= \|(d(u, a) - d(v, a))x\| = \\ &= |d(u, a) - d(v, a)| \leq d(u, v). \end{aligned}$$

Consequently

$$\begin{aligned} \left| \int f d\mu \right| &= |(f(a) - f(b) | x)| = \\ &= |(-d(b, a)x | x)| = d(a, b) \leq \|\mu\|_{MK}^*. \quad \square \end{aligned}$$

Notice that, if $a > 0$ and $v \in X$ are such that $\|v\| \leq a$, then $B_a(X, v) \stackrel{def}{=} B_a(X) \cap cabv(X, v) \neq \emptyset$ (because $\|\delta_t v\| = \|v\| \leq a$ for any $t \in T$).

On a non empty set $A \subset cabv(X)$, one can consider the following distances:

- The variational distance given via $d_{\|\cdot\|}(\mu, \nu) = \|\mu - \nu\|$.
- The Monge-Kantorovich distance given via $d_{MK}(\mu, \nu) = \|\mu - \nu\|_{MK}$.

- Assuming that $A - A \subset cabv(X, 0)$, the modified Monge-Kantorovich distance given via $d_{MK}^*(\mu, \nu) = \|\mu - \nu\|_{MK}^*$.

We shall mainly work in the particular case when $A = B_a(X, v)$ with $\|v\| \leq a$. On such $B_a(X, v)$ the last two distances are equivalent (Theorem 17): for μ and ν in $B_a(X, v)$, one has

$$d_{MK}(\mu, \nu) \leq d_{MK}^*(\mu, \nu) \leq d_{MK}(\mu, \nu)(\text{diam}(T) + 1).$$

In the next paragraph, we shall consider on such a set A another distance (namely the Hanin distance).

Before passing further, it is our duty to lay stress upon the fact that, maybe, a more honest name for the (modified) Monge-Kantorovich distance would have been Kantorovich-Rubinstein distance or Lipschitz distance.

Theorem 20. *Let $a > 0$ and $v \in X$ be such that $\|v\| \leq a$.*

1. *The (non empty) set $B_a(X, v)$ is weak* closed in $B_a(X)$, hence $B_a(X, v)$ is weak* compact. On $B_a(X, v)$, one has the equivalent metrics d_{MK} and d_{MK}^* .*

2. *For any $n \in \mathbb{N}$ (working for $X = K^n$), one has the supplementary result that $B_a(K^n, v)$, equipped with one of the equivalent metrics d_{MK} and d_{MK}^* is a compact, hence complete, metric space (its topology being equal to the weak* topology on $B_a(K^n, v)$).*

3. *In the particular case when $K = \mathbb{R}$, $n = 1$ and $v \geq 0$, one can consider the set $B_a^+(\mathbb{R}, v) = B_a(\mathbb{R}, v) \cap cabv^+(\mathbb{R})$, where $cabv^+(\mathbb{R}) = \{\mu \in cabv(\mathbb{R}) \mid \mu \geq 0\}$. Then $B_a^+(\mathbb{R}, v)$, equipped with one of the equivalent metrics d_{MK} and d_{MK}^* is a compact, hence complete, metric space (its topology being equal to the weak* topology on $B_a^+(\mathbb{R}, v)$).*

For $a = v = 1$, $B_1^+(\mathbb{R}, 1)$ is exactly the set of all probabilities on \mathcal{B} .

Proof. 1. We have $B_a(X, v) = B_a(X) \cap cabv(X, v)$. Because $B_a(X)$ is weak* compact, the result follows from Lemma 14.

2. The weak* topology of $B_a(K^n)$ coincides with the topology generated by the Monge-Kantorovich distance d_{MK} (Theorem 13). Hence, $B_a(K^n, v)$, being weak* closed in $B_a(K^n)$ which is weak* compact, will be also weak* compact. Therefore $B_a(K^n, v)$ is a compact subset of $B_a(K^n)$, considering on $B_a(K^n)$ the topology generated by the Monge-Kantorovich distance. Because the Monge-Kantorovich distance and the modified Monge-Kantorovich distance are equivalent on $B_a(K^n, v)$, it follows that $B_a(K^n, v)$ equipped either with the Monge-Kantorovich distance, or with the modified Monge-Kantorovich distance is a compact, hence complete, metric space.

3. In the particular case $K = \mathbb{R}$, $n = 1$, $v \geq 0$, all it remains to be proved is the fact that $cabv^+(\mathbb{R})$ is weak* closed. To this end, let $\mu \in cabv(\mathbb{R})$ be such that there exists $(\mu_\delta)_\delta$ net $cabv^+(\mathbb{R})$ with the property that $\mu_\delta \xrightarrow{\delta} \mu$ in $\mathcal{T}(w^*, \mathbb{R})$. This implies that for any $f \in C(\mathbb{R})$, $f \geq 0$, one has $\int f d\mu_\delta \xrightarrow{\delta} \int f d\mu$. Because $\int f d\mu_\delta \geq 0$ for any δ , it follows that $\int f d\mu \geq 0$. We succeeded in proving that for any $0 \leq f \in C(\mathbb{R})$, one has $\int f d\mu \geq 0$. So the functional $x'_\mu \in C(\mathbb{R})'$, given via $x'_\mu(f) = \int f d\mu$ is positive. The Riesz-Kakutani theorem says that this is equivalent to the fact that μ is positive, i.e. $\mu \in cabv^+(\mathbb{R})$. \square

The Hanin norm

The problem with the modified Monge-Kantorovich norm is the fact that it cannot be defined on the whole space $cabv(X)$. To be more precise, using the notation from Theorem 16, if one tries to extend p beyond $cabv(X, 0)$, one can obtain infinite values for the extension, as the following result shows.

Proposition 21. *Define $p : cabv(X) \rightarrow \overline{\mathbb{R}_+}$ via $p(\mu) = \sup\{|\int f d\mu| \mid f \in L_1(X)\}$. Then p is an extended seminorm (i.e. $p(\mu + \nu) \leq p(\mu) + p(\nu)$ and $p(\alpha\mu) = |\alpha|p(\mu)$ (with convention $0 \cdot \infty = 0$) for any $\mu, \nu \in cabv(X)$ and any $\alpha \in K$).*

We have the equivalence (for $\mu \in cabv(X)$): $p(\mu) < \infty \Leftrightarrow \mu \in cabv(X, 0)$.

Proof. The only fact which must be proved is the implication \Rightarrow in the enunciation. Let $\mu \in cabv(X)$ with $p(\mu) < \infty$. Accepting $\mu(T) \neq 0$, we shall arrive at a contradiction.

Indeed, let $x \in X$ with $\|x\| = 1$ and such that $(x \mid \mu(T)) = \|\mu(T)\| > 0$. Then, for any $n \in \mathbb{N}$, the function $f_n \in C(X)$ given via $f_n(t) = nx$ for any $t \in T$ is constant, so $\|f_n\|_L = 0$ and $\int f_n d\mu = n \|\mu(T)\| \xrightarrow{n} \infty$. Hence, because all $f_n \in L_1(X)$, it follows that $p(\mu) = \infty$ which is a contradiction. \square

L. Hanin ([11] and [10]) succeeded in "extending" the modified Monge-Kantorovich norm $\|\cdot\|_{MK}^*$ from $cabv(\mathbb{R}, 0)$ to the whole $cabv(\mathbb{R})$ (the "extension" is equivalent to $\|\cdot\|_{MK}^*$ on $cabv(\mathbb{R}, 0)$). Using approximately the same line of reasoning, we shall "extend" $\|\cdot\|_{MK}^*$ from $cabv(X, 0)$ to a norm $\|\cdot\|_H$ on $cabv(X)$, for an arbitrary Hilbert space X .

Define for any $\mu \in cabv(X)$:

$$\|\mu\|_H \stackrel{def}{=} \inf\{\|\nu\|_{MK}^* + \|\mu - \nu\| \mid \nu \in cabv(X, 0)\},$$

thus obtaining the map $\|\cdot\|_H : cabv(X) \rightarrow \mathbb{R}_+$.

Taking $\nu = 0$, we get $\|\mu\|_H \leq \|\mu\|$ for any $\mu \in cabv(X)$. If $\mu \in cabv(X, 0)$, taking $\nu = \mu$, we obtain $\|\mu\|_H \leq \|\mu\|_{MK}^*$.

Proposition 22. *For any $\mu \in cabv(X, 0)$, one has*

$$\|\mu\|_{MK}^* \leq \max(\text{diam}(T), 1) \|\mu\|_H.$$

Proof. For any $\nu \in cabv(X, 0)$, using Theorem 18 we have:

$$\begin{aligned} \|\mu\|_{MK}^* &= \|\nu + \mu - \nu\|_{MK}^* \leq \|\nu\|_{MK}^* + \|\mu - \nu\|_{MK}^* \leq \\ &\leq \|\nu\|_{MK}^* + \|\mu - \nu\| \text{diam}(T) = \\ &= \|\nu\|_{MK}^* + \|\mu - \nu\| + \|\mu - \nu\| (\text{diam}(T) - 1). \end{aligned}$$

In case $\text{diam}(T) \leq 1$, we get, for any $\nu \in cabv(X, 0)$: $\|\mu\|_{MK}^* \leq \|\nu\|_{MK}^* + \|\mu - \nu\|$ and passing to infimum we get $\|\mu\|_{MK}^* \leq \|\mu\|_H$.

In case $\text{diam}(T) > 1$, we get, for any $\nu \in cabv(X, 0)$:

$$\begin{aligned} \|\mu\|_{MK}^* &\leq \|\nu\|_{MK}^* + \|\mu - \nu\| + (\|\nu\|_{MK}^* + \|\mu - \nu\|)(\text{diam}(T) - 1) = \\ &= (\|\nu\|_{MK}^* + \|\mu - \nu\|)(1 + \text{diam}(T) - 1) \end{aligned}$$

and passing to infimum, we get $\|\mu\|_{MK}^* \leq \|\mu\|_H \text{diam}(T)$. \square

Before passing further, let us notice that, for any $f \in L(X)$ and any $\mu \in cabv(X)$, one has

$$\left| \int f d\mu \right| \leq \|\mu\|_H \|f\|_{BL}. \quad (11)$$

Indeed, for any $\nu \in cabv(X, 0)$ we have:

$$\begin{aligned} \left| \int f d\mu \right| &= \left| \int f d(\mu - \nu) + \int f d\nu \right| \leq \left| \int f d\nu \right| + \left| \int f d(\mu - \nu) \right| \leq \\ &\leq \|\nu\|_{MK}^* \|f\|_L + \|\mu - \nu\| \|f\| \leq \\ &\leq \|f\|_{BL} (\|\nu\|_{MK}^* + \|\mu - \nu\|), \end{aligned}$$

with (10). Due to the arbitrariness of ν , (11) is proved.

Theorem 23. 1. The functional $\|\cdot\|_H : cabv(X) \rightarrow \mathbb{R}_+$ is a norm on $cabv(X)$ which generates a topology weaker than the variational topology generated by $\|\cdot\|$: $\|\mu\|_H \leq \|\mu\|$ for any $\mu \in cabv(X)$.

2. On $cabv(X, 0)$, the modified Monge-Kantorovich norm $\|\cdot\|_{MK}^*$ and the restriction of $\|\cdot\|_H$ are equivalent:

$$\|\mu\|_H \leq \|\mu\|_{MK}^* \leq \max(\text{diam}(T), 1) \|\mu\|_H$$

(hence $\|\mu\|_H = \|\mu\|_{MK}^*$, if $\text{diam}(T) \leq 1$) for any $\mu \in cabv(X, 0)$.

Proof. It remains to be proved that $\|\cdot\|_H$ is a norm on $cabv(X, 0)$. First we prove that $\|\cdot\|_H$ is a seminorm.

Because $cabv(X, 0) = cabv(X, 0) + cabv(X, 0)$, we have for any μ_1 and μ_2 in $cabv(X)$:

$$\|\mu_1 + \mu_2\|_H = \inf\{\|\nu_1 + \nu_2\|_{MK}^* + \|\mu_1 + \mu_2 - \nu_1 - \nu_2\| \mid \nu_1, \nu_2 \in cabv(X, 0)\}.$$

Because

$$\begin{aligned} & \|\nu_1 + \nu_2\|_{MK}^* + \|\mu_1 + \mu_2 - \nu_1 - \nu_2\| \leq \\ & \leq \|\nu_1\|_{MK}^* + \|\nu_2\|_{MK}^* + \|\mu_1 - \nu_1\| + \|\mu_2 - \nu_2\|, \end{aligned}$$

for any $\nu_1, \nu_2 \in cabv(X, 0)$, we pass to infimum obtaining $\|\mu_1 + \mu_2\|_H \leq \|\mu_1\|_H + \|\mu_2\|_H$.

For $\alpha \in K$ and $\mu \in cabv(X)$, one has $\|\alpha\mu\|_H = |\alpha| \|\mu\|_H$. This is obvious for $\alpha = 0$. If $\alpha \neq 0$, using the equality $\alpha cabv(X, 0) = \{\alpha\mu \mid \mu \in cabv(X, 0)\} = cabv(X, 0)$ we have:

$$\begin{aligned} \|\alpha\mu\|_H &= \inf\{\|\nu\|_{MK}^* + \|\alpha\mu - \nu\| \mid \nu \in cabv(X, 0)\} = \\ &= \inf\{\|\alpha\nu_1\|_{MK}^* + \|\alpha\mu - \alpha\nu_1\| \mid \nu_1 \in cabv(X, 0)\} = \\ &= |\alpha| \inf\{\|\nu_1\|_{MK}^* + \|\mu - \nu_1\| \mid \nu_1 \in cabv(X, 0)\} = |\alpha| \|\mu\|_H. \end{aligned}$$

Finally, we show that, for $\mu \in cabv(X)$, one has the implication: $\|\mu\|_H = 0 \Rightarrow \mu = 0$. Indeed, if $\|\mu\|_H = 0$, we have $\int f d\mu = 0$ for any $f \in L(X)$, using (11). This implies $\int f d\mu = 0$ for any $f \in BL_1(X)$, hence $\|\mu\|_{MK} = 0$, i.e. $\mu = 0$. \square

Corollary 24. On $cabv(X, 0)$, the restrictions of $\|\cdot\|_H$ and $\|\cdot\|_{MK}$ are equivalent: for any $\mu \in cabv(X, 0)$ one has

$$\frac{1}{\max(\text{diam}(T), 1)} \|\mu\|_{MK} \leq \|\mu\|_H \leq (\text{diam}(T) + 1) \|\mu\|_{MK},$$

hence, if $\text{diam}(T) \leq 1$, we hence

$$\|\mu\|_{MK} \leq \|\mu\|_H \leq 2 \|\mu\|_{MK}.$$

Proof. Using Theorems 18 and 23, we get for $\mu \in \text{cabv}(X, 0)$:

$$\|\mu\|_H \leq \|\mu\|_{MK}^* \leq (\text{diam}(T) + 1) \|\mu\|_{MK}$$

and

$$\|\mu\|_H \geq \frac{1}{\max(\text{diam}(T), 1)} \|\mu\|_{MK}^* \geq \frac{1}{\max(\text{diam}(T), 1)} \|\mu\|_{MK}. \quad \square$$

Definition 25. The norm $\|\cdot\|_H$ on $\text{cabv}(X)$ will be called *the Hanin norm*.

As usual, we present the afferent notations. The topology on $\text{cabv}(X)$, generated by $\|\cdot\|_H$, will be called *the Hanin topology* and it will be denoted by $\mathcal{T}(H, X)$. This topology induces the topology $\mathcal{T}(H, X, a)$ on $B_a(X)$. For a sequence $(\mu_n)_n \subset \text{cabv}(X)$ and for $\mu \in \text{cabv}(X)$, we write $\mu_n \xrightarrow[n]{H} \mu$ to denote the fact that $(\mu_n)_n$ converges to μ in the topology $\mathcal{T}(H, X)$. Finally, on any $\emptyset \neq A \subset \text{cabv}(X)$, $\|\cdot\|_H$ generates *the Hanin metric* d_H given via $d_H(\mu, \nu) = \|\mu - \nu\|_H$ for any μ, ν in A .

Theorem 23 is very important. Due to the equivalence of $\|\cdot\|_H$ and $\|\cdot\|_{MK}^*$ on $\text{cabv}(X, 0)$, it follows that the metrics d_H and d_{MK}^* are equivalent on any $B_a(X, v)$. Hence, in the enunciation of Theorem 20, one can add d_H to the previous equivalent metrics d_{MK} and d_{MK}^* . Consequently we have:

Corollary 26. *If $a > 0$ and $v \in X$ are such that $\|v\| \leq a$, then the metrics d_{MK} , d_{MK}^* and d_H are equivalent on the non empty set $B_a(X, v)$.*

Consequently, in the enunciation of Theorem 20, one can replace (three times) the expression "the equivalent metrics d_{MK} and d_{MK}^ " with the expression "the equivalent metrics d_{MK} , d_{MK}^* and d_H ", thus augmenting the enunciation.*

Theorem 27. *Let a and b in T , $a \neq b$ and $x \in X$, $\|x\| = 1$. Then*

$$\|\delta_a x\|_H = 1$$

and

$$\frac{1}{\max(\text{diam}(T), 1)} d(a, b) \leq \|\delta_a x - \delta_b x\|_H \leq d(a, b).$$

Hence, in case $\text{diam}(T) \leq 1$, $\|\delta_a x - \delta_b x\|_H = d(a, b)$. For $X = K$ and $x = 1$, we have

$$\frac{1}{\max(\text{diam}(T), 1)} d(a, b) \leq \|\delta_a - \delta_b\|_H \leq d(a, b).$$

Proof. Again define $f = \varphi_x$, where $\varphi_x : T \rightarrow X$, $\varphi_x(t) = x$ for any $t \in T$ and notice that $\|f\|_{BL} = \|f\| = 1$.

For $\mu = \delta_a x$ use (11) and get

$$\left| \int f d\mu \right| = (x \mid \mu(T)) = (x \mid x) = 1 \leq \|\mu\|_H \leq \|\mu\| = 1.$$

Now, writing $\nu = \delta_a x - \delta_b x \in \text{cabv}(X, 0)$ and using Theorems 19 and 23, we get:

$$\|\nu\|_H \leq \|\nu\|_{MK}^* = d(a, b)$$

and

$$\|\nu\|_H \geq \frac{1}{\max(\text{diam}(T), 1)} \|\nu\|_{MK}^* = \frac{1}{\max(\text{diam}(T), 1)} d(a, b). \quad \square$$

Infinite dimensional extensions do not work

In this subparagraph we present a counterexample showing that neither Theorem 11, nor Theorem 12, can be extended for general Hilbert spaces instead of K^n .

Counterexample. Our compact metric space (T, d) will be given as follows: $T = \{1, 2\}$ with metric $d(i, j) = 1$ if $i \neq j$ and $d(i, j) = 0$ if $i = j$. Hence, on T we have the discrete topology and the Borel sets are $\mathcal{B} = \mathcal{P}(T)$. Our Hilbert space will be l^2 .

Any function $f : T \rightarrow l^2$ is continuous, even Lipschitz, and simple. We identify such a function f giving $f(1) = (a_{1m})_m$ and $f(2) = (a_{2m})_m$. Clearly $\|f\|_L = \|f(1) - f(2)\|$.

A measure $\mu \in \text{cabv}(l^2)$ is identified giving $\mu(\{1\}) = (b_{1m})_m \in l^2$ and $\mu(\{2\}) = (b_{2m})_m \in l^2$. Hence, the total variation $|\mu|(T) = \|\mu\| = \|\mu(\{1\})\| + \|\mu(\{2\})\|$.

Using the previous notations, we have:

$$\int f d\mu = (f(1) \mid \mu(\{1\})) + (f(2) \mid \mu(\{2\})) = \sum_{m=1}^{\infty} (a_{1m} \overline{b_{1m}} + a_{2m} \overline{b_{2m}}).$$

We shall be concerned with the case when $\mu \in cabv(l^2, 0)$, which means $\mu(\{1\}) + \mu(\{2\}) = 0$, i.e. $(b_{1m} + b_{2m})_m = 0$, hence $b_{2m} = -b_{1m}$ for any m . In this case, we shall identify $\mu \equiv b = (b_m)_m \in l^2$, where $b_{1m} = b_m$ and $b_{2m} = -b_m$ for any m .

We have $\|\mu\| = \|b\| + \|b\| = 2\|b\|$ and $\int f d\mu = \sum_{m=1}^{\infty} (a_{1m} - a_{2m}) \overline{b_m}$.

Writing $a = (a_{1m} - a_{2m})_m \in l^2$, we notice that $\int f d\mu = (a \mid b)$. Notice also that $f \in L_1(l^2)$ means $\|f(1) - f(2)\| \leq 1$ i.e. $\|a\| \leq 1$.

The final preliminary fact is that for $b \equiv \mu \in cabv(l^2, 0)$, one has $\|\mu\|_{MK}^* = \|b\|$. Indeed $\|\mu\|_{MK}^* = \sup\{|\int f d\mu| \mid \|f\|_L \leq 1\}$. For f identified as above, we saw that $\|f\|_L \leq 1$ means $\|a\| \leq 1$, hence $|\int f d\mu| = |(a \mid b)| \leq \|b\|$, consequently $\|\mu\|_{MK}^* \leq \|b\|$. On the other hand, let us take $a = (a_m)_m \in l^2$ such that $\|a\| = 1$ and $(a \mid b) = \|b\|$. Define $a_{1m} = a_m$ and $a_{2m} = 0$, for every $m \in \mathbb{N}$. We got the function $f : T \rightarrow l^2$ identified as above: $f(1) = (a_{1m})_m = (a_m)_m$ and $f(2) = (a_{2m})_m = 0$. Then, for this f one has $\|f(1) - f(2)\| = \|a\| = 1$ (hence $f \in L_1(l^2)$) and $\int f d\mu = (a \mid b) = \|b\|$. It follows that $\|b\| \leq \|\mu\|_{MK}^*$ and the equality $\|\mu\|_{MK}^* = \|b\|$ is proved.

Practically, we proved the existence of the linear and isometric isomorphism $(cabv(l^2, 0), \|\cdot\|_{MK}^*) \equiv l^2$, via $\mu \equiv b$ as above.

Let us consider an arbitrary bounded sequence $(b^m)_m$ in l^2 and let $a > 0$ be such that $\|b^m\| \leq \frac{a}{2}$ for any m . According to the previous isomorphism considerations, identify each $b^m \equiv \mu^m \in cabv(l^2, 0)$. We saw that $\|\mu^m\| = 2\|b^m\| \leq a$, hence $\mu^m \in B_a(l^2)$ for any m .

According to Theorem 18 and previous considerations we have the equivalences: $\mu^m \xrightarrow[m]{MK} 0$ if and only if $\mu^m \xrightarrow[m]{MK^*} 0$ if and only if $b^m \xrightarrow[m]{} 0$, the last convergence being in l^2 .

On the other hand, we have the equivalences: $\mu^m \xrightarrow[m]{w^*} 0$ if and only if $(a \mid b^m) \xrightarrow[m]{} 0$ for any $a \in l^2$ if and only if $b^m \xrightarrow[m]{w^*} 0$ (the last convergence being weak convergence in l^2). We must prove the equivalence: $\mu^m \xrightarrow[m]{w^*} 0 \Leftrightarrow (a \mid b^m) \xrightarrow[m]{} 0$ for any $a \in l^2$. To prove " \Rightarrow ", take $a = (a_n)_n \in l^2$ and

define $f \in C(l^2)$ identified via $f(1) = (a_n)_n$ and $f(2) = 0$. Then $(a \mid b^m) = \int f d\mu_m^m \rightarrow 0$. To prove " \Leftarrow ", take $f \in C(l^2)$ identified via $f(1) = (a_{1n})_n$ and $f(2) = (a_{2n})_n$ and define $a = (a_{1n} - a_{2n})_n \in l^2$. Then $\int f d\mu_m^m = (a \mid b^m) \rightarrow 0$.

We arrived at the end of our discussion. Accept that the assertion in Theorem 12 is valid for the separable Hilbert space l^2 instead of K^n for any number $a > 0$. This means to accept the following assertion (taking $\mu^m - \mu$ instead of μ^m and 0 instead of μ in Theorem 12): For any number $a > 0$ and for any sequence $(\mu^m)_m \subset B_a(l^2, 0)$, one has the equivalence: $\mu^m \xrightarrow[m]{MK^*} 0$ if and only if $\mu^m \xrightarrow[m]{w^*} 0$. This last assertion "translated" in view of the previous considerations means: For any number $a > 0$ and for any bounded sequence $(b^m)_m \subset l^2$ such that $\|b^m\| \leq \frac{a}{2}$ for any m , one has the equivalence: $b^m \xrightarrow[m]{} 0$ (in l^2) if and only if $b^m \xrightarrow[m]{} 0$ (weakly in l^2).

The last equivalence is clearly false. For a concrete example of failure, one can take $a = \frac{2\pi}{\sqrt{6}}$ and the sequence $(b^p)_p$ with $b^p = (b_n^p)_n$, as follows:

For $p = 1$: $b_n^1 = \frac{1}{n}$, $n \in \mathbb{N}$.

For $p > 1$: $b_n^p = \begin{cases} 0, & \text{if } n < p \\ \frac{1}{n-p+1}, & \text{if } n \geq p \end{cases}$.

Then $\|b^p\| = \frac{\pi}{\sqrt{6}} = \frac{a}{2}$ for any p . Clearly the assertion $b^p \xrightarrow[p]{} 0$ in l^2 is false.

On the other hand, let us take $a = (a_n)_n \in l^2$. Defining, for any $p \in \mathbb{N}$, the new sequence $a(p) = (a_{n+p-1})_n$, one has $\|a(p)\|^2 = \sum_{n=1}^{\infty} |a_{n+p-1}|^2 \xrightarrow[p]{} 0$. For any p , one can see that $(a \mid b^p) = (a(p) \mid b^1)$. It follows that $|(a \mid b^p)| \leq \|a(p)\| \|b^1\| \xrightarrow[p]{} 0$ and this shows that $b^p \xrightarrow[p]{} 0$ weakly in l^2 .

Remark. The previous counterexample shows that the "extension" of Theorem 12 for an arbitrary X instead of K^n is false. At the same time, the counterexample shows that the "extension" of Theorem 11 for an arbitrary X instead of K^n is false too. Indeed, accepting for instance that $BL_1(l^2)$ is relatively compact in $C(l^2)$, we can repeat the "proof" of Theorem 12 for l^2 instead of K^n and arrive at the false conclusion that weak* convergence and Monge-Kantorovich convergence are the same in $cabv(l^2)$.

New metrics on T generated by the previous norms on $cabv(X, 0)$

In this last subparagraph, we shall discuss the new metrics on T which are

generated by the norms on $cabv(X, 0)$ which have been introduced throughout the paper.

To begin, let us choose an arbitrary $x \in X$ with $\|x\| = 1$ which will be fixed from now on (in the special case $X = K$, we take canonically $x = 1$). Recall that $\|\delta_t x\| = \|\delta_t x\|_{MK} = \|\delta_t x\|_H = 1$ (relation (3) and Theorem 27). We define the injective map $V : T \rightarrow cabv(X)$ via $V(t) = \delta_t x$. Then $\delta_t x - \delta_s x \in cabv(X, 0)$ and, for any norm p on $cabv(X, 0)$, one obtains the metric ρ_p on T given via $\rho_p(t, s) = p(\delta_t x - \delta_s x)$.

The fact that $\rho_{\|\cdot\|}(t, s) = 2$ for $t \neq s$ shows that the metric $\rho_{\|\cdot\|}$ generates the discrete topology on T , being metrically insensitive (rigid) (see relation (4)).

To complete our discussion, we shall need.

Lemma 28. *For any $a \neq b$ in T , there exists $g \in L(\mathbb{R})$ having the following properties:*

i) $g(a) = 0$, $g(b) = 1$ and $0 \leq g(t) \leq 1$ for any $t \in T$.

ii) $\|g\| = 1$, $\|g\|_L \leq \frac{1}{d(a,b)}$, hence $\|g\|_{BL} \leq \frac{1+d(a,b)}{d(a,b)}$.

Proof. Let us define $g : T \rightarrow \mathbb{R}$, via

$$g(t) = \frac{d(t, a)}{d(t, a) + d(t, b)}.$$

Then i) and $\|g\| = 1$ follow immediately. It remains to be proved that $\|g\|_L \leq \frac{1}{d(a,b)}$.

Indeed, for any x and y in T , one has

$$\begin{aligned} |g(x) - g(y)| &= \frac{|d(x, a)d(y, b) - d(x, b)d(y, a)|}{(d(x, a) + d(x, b))(d(y, a) + d(y, b))} \leq \\ &\leq \frac{|d(x, a)d(y, b) - d(x, b)d(y, a)|}{d(a, b)(d(y, a) + d(y, b))} = \\ &= \frac{|d(x, a)d(y, b) - d(y, a)d(y, b) + d(y, a)d(y, b) - d(y, a)d(x, b)|}{d(a, b)(d(y, a) + d(y, b))} \leq \\ &\leq \frac{d(y, b)|d(x, a) - d(y, a)| + d(y, a)|d(x, b) - d(y, b)|}{d(a, b)(d(y, a) + d(y, b))} \leq \\ &\leq \frac{d(y, b)d(x, y) + d(y, a)d(x, y)}{d(a, b)(d(y, a) + d(y, b))} \leq \frac{d(x, y)}{d(a, b)}. \quad \square \end{aligned}$$

Theorem 29. *The metrics d , $\rho_{\|\cdot\|_{MK}}$, $\rho_{\|\cdot\|_{MK}^*}$ and $\rho_{\|\cdot\|_H}$ are Lipschitz equivalent. Namely, one has*

$$\frac{1}{1 + \text{diam}(T)} d \leq \rho_{\|\cdot\|_{MK}} \leq d, \quad (12)$$

$$\rho_{\|\cdot\|_{MK}^*} = d \quad (13)$$

and

$$\frac{1}{\max(\text{diam}(T), 1)} d \leq \rho_{\|\cdot\|_H} \leq d. \quad (14)$$

The metric $\rho_{\|\cdot\|}$ generates the discrete topology on T , being equivalent to d if and only if T is finite.

Proof. Relation (13) is proved in Theorem 19 and relation (14) is proved in Theorem 27. Because $\rho_{\|\cdot\|_{MK}} \leq d$ (see relation (5)), all it remains to be proved is the fact (completely proving (12)) that

$$\frac{1}{1 + \text{diam}(T)} d \leq \rho_{\|\cdot\|_{MK}}. \quad (15)$$

To this end, for the previous $x \in X$ with $\|x\| = 1$, define $f = gx$, where g is the function from Lemma 28, for arbitrary $a \neq b$ in T . Then, for any $\mu \in \text{cabv}(X)$ one has

$$\begin{aligned} \left| \int f d\mu \right| &\leq \|\mu\|_{MK} \|f\|_{BL} = \|\mu\|_{MK} \|g\|_{BL} \leq \\ &\leq \|\mu\|_{MK} \frac{1 + d(a, b)}{d(a, b)} \leq \frac{1 + \text{diam}(T)}{d(a, b)} \|\mu\|_{MK} \end{aligned}$$

and this implies

$$\|\mu\|_{MK} \geq \frac{d(a, b)}{1 + \text{diam}(T)} \left| \int f d\mu \right|.$$

Taking $\mu = \delta_a x - \delta_b x$, one has $\int f d\mu = ((f(a) - f(b)) \mid x) = -(x \mid x) = -1$, hence $\|\mu\|_{MK} = \rho_{\|\cdot\|_{MK}}(a, b) \geq \frac{d(a, b)}{1 + \text{diam}(T)}$ and (15) is proved. \square

Remarks. 1. The equivalence between the metrics d and the metrics $\rho_{\|\cdot\|_{MK}}$, $\rho_{\|\cdot\|_{MK}^*}$ and $\rho_{\|\cdot\|_H}$ can be obtained immediately observing the fact that (T, d) is a metric compact space and the other metrics generate weaker topologies. Of course, the result of Theorem 29 is stronger.

2. Because (see Lemma 28) one has, for any $a \neq b$ in T , the inequality $\|g\|_{BL} \leq \frac{1+d(a,b)}{d(a,b)}$, the proof of Theorem 29 shows that, actually, for any a and b in T : $\frac{1}{1+diam(T)}d(a,b) \leq \frac{1}{1+d(a,b)}d(a,b) \leq \rho_{\|\cdot\|_{MK}}(a,b) \leq d(a,b)$ thus improving (12).

3. In some cases, the evaluations given in Theorem 29 can be improved. For instance, let us consider a number $A > 0$, obtaining the compact metric space $T = [0, A]$ with metric d given by the natural metric of \mathbb{R} . For any a and b in T , we use, for the computation of $\|\delta_a - \delta_b\|_{MK} = \rho_{\|\cdot\|_{MK}}(a,b)$, affine functions $f : T \rightarrow \mathbb{R}$ of the form $f(x) = mx + n$ and, with some effort, we can conclude that $\rho_{\|\cdot\|_{MK}}(a,b) \geq \frac{2|a-b|}{2+A}$. This means that, in this case, one has

$$\frac{2}{2 + diam(T)}d(a,b) \leq \rho_{\|\cdot\|_{MK}}(a,b). \quad (16)$$

Relation (16) improves relation (15), i.e. improves relation (12).

At the same time, taking A arbitrary small, one can conclude from (16) that, for any $\varepsilon \in (0, 1)$, there exists a compact metric space (T, d) such that $(1 - \varepsilon)d \leq \rho_{\|\cdot\|_{MK}} \leq d$. Hence, it is natural to try to solve the following

Open Problem: Find new evaluations, sharper than (12), (13) and (14).

References

- [1] P. Appel, Mémoire sur les déblais et les remblais des systèmes continus ou discontinus. Mémoires présentées par divers Savants à l'Académie des Sciences de l'Institut de France, Paris 29 (1887), 1-208.
- [2] P. Billingsley, Convergence of Probability Measures, second ed., John Wiley & Sons, New York, 1999.
- [3] I. Chişescu, Function Spaces, Ed. Şt. Encicl. Bucureşti, 1983 (in Romanian).
- [4] J. Diestel, J.J. Uhl, Jr., Vector Measures, Math. Surveys, vol. 15, Amer. Math. Soc., Providence, R.I., 1977.
- [5] N. Dinculeanu, Vector Measures, VEB Deutscher Verlag der Wissenschaften, Berlin, 1966.
- [6] R.M. Dudley, Real Analysis and Probability, second ed., Cambridge University Press, Cambridge, 2002.
- [7] J. Dugundji, Topology, fourth ed., Allyn and Bacon, Boston, 1968.
- [8] N. Dunford, J.T. Schwartz, Linear Operators, Part I: General Theory, Interscience Publishers, New York, 1957.

- [9] G.B. Folland, *Real Analysis. Modern Techniques and Their Applications*, second ed., John Wiley & Sons, New York, 1999.
- [10] L.G. Hanin, Kantorovich-Rubinstein norm and its application in the theory of Lipschitz spaces, *Proc. Amer. Math. Soc.* 115 (1992) 345-352.
- [11] L.G. Hanin, An extension of the Kantorovich norm. In *Monge-Ampère Equation: Applications to geometry and optimization* (Deerfield Beach, FL, 1997), vol. 226 of *Contemp. Math.*, Amer. Math. Soc., Providence, RI (1999), 113-130.
- [12] L.V. Kantorovich, On the translocation of masses, *C.R. (Doklady) Acad. Sci. URSS* 37 (1942) 199-201.
- [13] L.V. Kantorovich, G.P. Akilov, *Functional Analysis*, second ed., Pergamon Press, New York, 1982.
- [14] L.V. Kantorovich, G.S. Rubinstein, On a certain function space and some extremal problems, *C.R. (Doklady) Acad. Sci. URSS* 115 (1957) 1058-1061 (in Russian).
- [15] L.V. Kantorovich, G.S. Rubinstein, On a space of completely additive functions, *Vestnik Leningrad Univ.* 13 (1958) 52-59 (in Russian).
- [16] J.L. Kelley, *General topology*, D. Van Nostrand Company, Toronto-New York-London, 1955.
- [17] J. Lukeš, J. Maly, *Measure and Integral*, Matfyzpress, Publishing House of the Faculty of Mathematics and Physics, Charles University, Prague, 1995.
- [18] G. Monge, *Mémoire sur la théorie des déblais et des remblais*. In *Histoire de l'Académie Royale des Sciences de Paris*, 1781, 666-704.
- [19] K.R. Parthasarathy, *Probability measures on metric spaces*, Academic Press, New York, 1967.
- [20] H.H. Schaefer, *Topological vector spaces*, third printing corrected, Springer, New York, 1971.
- [21] C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, vol. 58, Amer. Math. Soc., Providence, R.I., 2003.